

STIC Biotechnology Systems Branch**RAW SEQUENCE LISTING**
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/505,328A
Source: Pat
Date Processed by STIC: 4/28/06

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENT/IN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER** **VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06



PCT

RAW SEQUENCE LISTING

DATE: 04/28/2006

PATENT APPLICATION: US/10/505,328A

TIME: 09:38:57

Input Set : A:\Sequence.txt

Output Set: Y:\CRF4\04282006\J505328A.raw

C--> 2 <110> APPLICANT: Korea Advanced Institute of Science and Technology
4 <120> TITLE OF INVENTION: CONSTRUCTION OF NOVEL STRAINS CONTAINING MINIMIZING
5 GENOME BY Tn5-COUPLED Cre/loxP EXCISION SYSTEM
7 <130> FILE REFERENCE: 02730.0020.PCUS00
9 <140> CURRENT APPLICATION NUMBER: 10/505,328A
11 <141> CURRENT FILING DATE: 2004-08-23
11 <150> PRIOR APPLICATION NUMBER: PCT/KR02/02033 -
12 <151> PRIOR FILING DATE: 2002-10-31
14 <150> PRIOR APPLICATION NUMBER: KR 10-2002-0009647
15 <151> PRIOR FILING DATE: 2002-02-22
17 <160> NUMBER OF SEQ ID NOS: 13
19 <170> SOFTWARE: KopatentIn 1.71
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 2437
23 <212> TYPE: DNA
24 <213> ORGANISM: Artificial Sequence
26 <220> FEATURE:
27 <223> OTHER INFORMATION: chemically synthesized TnKGloXP
30 <400> SEQUENCE: 1
31 attcaggctg cgcaactggt gggaagggcg atcgggtgagg gcctcttcgc tattaagcca 60
33 gctgtctctt atacacatct caaccatcat cgatgaattc gagctcggta cccgggttga 120
35 actgcggatc ttgcggccgc aaaaattaaa aatgaagttt tgacgggtatc gaaccccaga 180
37 gtcccgctca gaagaactcg tcaagaaggc gatagaaggc gatgcgctgc gaatcgggag 240
39 cggcgatacc gtaaagcacg aggaagcggc cagcccattc gccgccaagc tcttcagcaa 300
41 tatcacgggt agccaacgct atgtcctgat agcgggtccgc cacacccagc cggccacagt 360
43 cgatgaatcc agaaaagcgg ccattttcca ccatgatatt cggcaagcag gcacgcacct 420
45 ggggtcacgac gagatcctcg ccgtcgggca tccgcgcctt gagcctggcg aacagttcgg 480
47 ctggcgcgag cccctgatgc tcttcgtcca gatcctctg atcgacaaga ccggcttcca 540
49 tccgagtagc tgcctcgctc atgcgatgtt tcgcttggtg gtcgaatggg caggtagccg 600
51 gatcaagcgt atgcagccgc cgcattgcat cagccatgat ggatactttc tcggcaggag 660
53 caaggtgaga tgacaggaga tccctgcccc gcacttcgac caatagcagc cagtccttcc 720
55 ccgcttcagt gacaacgtcg agcacagctg cgcaaggaaac gcccgctcgt gccagccacg 780
57 atagcccgcg tgcctcgctc tggagttcat tcaggggcacc ggacaggctc gtcttgacaa 840
59 aaagaaccgg gcgcccctgc gctgacagcc ggaacacggc ggcatcagag cagccgattg 900
61 tctgttgtgc ccagtcatac ccgaatagcc tctccaccca agcggccgga gaacctgcgt 960
63 gcaatccatc ttgttcaatc atgcgaaacg atcctcatcc tgtctcttga tccactagat 1020
65 tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 1080
67 aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgcacgatg 1140
69 aattgatccg aagttcctat tctctagaaa gtataggaaac ttcgaattgt cgacaagctt 1200
71 gatctggctt atcgaaatta atacgactca ctatagggag accggaattc attatttgta 1260
73 gagctcatcc atgccatgtg taatcccagc agcagttaca aactcaagaa ggaccatgtg 1320
75 gtcacgcttt tcggtgggat ctttcgaaag ggcagattgt gtcgacaggc aatggttgtc 1380
77 tggtaaaagg acaggggcat cgccaattgg agtattttgt tgataatggc ctgctagttg 1440

(pg.5)

file://C:\CRF4\Outhold\VsrrJ505328A.htm

4/28/2006

RAW SEQUENCE LISTING

DATE: 04/28/2006

PATENT APPLICATION: US/10/505,328A

TIME: 09:38:57

Input Set : A:\Sequence.txt

Output Set: N:\CRF4\04282006\J505328A.raw

```
79 aacggatcca tcttcaatgt tgtggcgaat tttgaagtta gctttgatto cattcttttg 1500
81 tttgtctgcc gtgatgtata cattgtgtga gttatagttg tactcgagtt tgtgtccgag 1560
83 aatgtttcca tcttctttaa aatcaatacc ttttaactcg atacgattaa caagggatc 1620
85 accttcaaac ttgacttcag cagcgctctt gtagttcccg tcatccttga aagatatag 1680
87 gcgttctctg acataacctt cgggcatggc actcttgaaa aagtcatgcc gtttcatatg 1740
89 atccggataa cgggaaaagc attgaacacc ataagagaaa gtagtgacaa gtgttggcca 1800
91 tggaacaggt agttttccag tagtgcaaat aaatttaagg gtaagttttc cgtatgttgc 1860
93 atcaccttca cctctccac tgacagaaaa tttgtgcca ttaacatcac catctaattc 1920
95 aacaagaatt gggacaactc cagtgaagaag ttcttctcct ttaactatct tttctaccgg 1980
97 taccggggga tctctagag tcgacctgca ggcattgcaag cttggcgtaa tcatggtc 2040
99 agctgtttcc tgtgtgaat tgttarccgc tcacaattcc acacaacata cgaagcggaa 2100
101 gcataaagtg taaagcctgg ggtgcctaat gagttagcta actcacatta attgcgttgc 2160
103 gctcactgcc cgttttccag tcgggaaatc caagggcgaa ttcgagctcg gtaccggggc 2220
105 cccctcagag ggacctata acttcgtata gcatacatta tacgaagtta tattaaggg 2280
107 tccggatcct ctagagtaga cctctagagt cgacctgcag gcattgcaagc ttcagggttg 2340
109 agatgtgtat aagagacagc tgcatctaag aatcggccaa cgcgcgggga gaggcggtt 2400
111 gcgtattggg cgtctctcgc ctctctcgt cactgac 2437
114 <210> SEQ ID NO: 2
115 <211> LENGTH: 1511
116 <212> TYPE: DNA
117 <213> ORGANISM: Artificial Sequence
119 <220> FEATURE:
120 <223> OTHER INFORMATION: chemically synthesized TnClopX
123 <400> SEQUENCE: 2
124 attcaggctg cgaactgtt ggggaagggc atcgggtgagg gcctcttctgc tattaagcca 60
126 gctgtctctt atcacatct caaccatcat cgatgaattc gagctcggta ccgcaaaaaa 120
128 taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac 180
130 caatgcttaa tcagtgggc accaataact gccttaaaaa aattacgccc cgcctggcca 240
132 ctcatcgacg tactgttgta attcattaag cattctgcog acatggaage catcacagac 300
134 ggcattgatg acctgaatcg ccagcggcat cagcaccttg tcgccttgcg tataatattt 360
136 gccatgggtg aaaaacgggg cgaagaagtt gtccatattg gcaacgttta aatcaaaact 420
138 ggtgaaactc acccagggat tggctgagac gaaaaacata ttctcaataa accctttagg 480
140 gaaataggcc aggttttcac cgtaacacgc cacatcttgc gaatatatgt gtagaaactg 540
142 cgggaaatcg tcgtggtatt cactccagag cgatgaanaa gtttcagttt gtcattggaa 600
144 aacgggtgta caagggtgaa cactatccca tatcaccagc tcaccgtctt tcattgocat 660
146 aoggaatttc ggtgagcat tcctcaggcg ggcaagaatg tgaataaagg ccggataaaa 720
148 cttgtgctta tttttcttta cggctcttta aaaggccgta atatccagct gaacgggtotg 780
150 gttataggta cattgagcaa ctgactgaaa tgcctcaaaa tgttctttac gatgccattg 840
152 ggatatatca acggtggtat atccagtgat ttttttctcc attttagctt ccttagctcc 900
154 tgaatatctc gataactcaa aaaatacggc cggtagtgat cttatttcat tatggtgaaa 960
156 gttggaacct cttacgtgcc gatcaacgtc tcattttcgc caaaagtttg cccagggctt 1020
158 cccggatcca acagggacac caggatttat ttattctgcg aagtgatctt ccgtcacagg 1080
160 tattttattc gcgcaaagtg cgtcgggtga tgcgcgaac ttaactgatt agtgtatgat 1140
162 ggtgtttttg aggtgctcca gtggtctctg ttctctatcag catcgatgaa ttgatcogaa 1200
164 gttcttattc tctagaaagt ataggaactt cgaattgtcg acaagcttga tctggcttat 1260
166 cgaataaatc acgactcact atagggagac cgggaattcg gctcggtaac gggccccccc 1320
168 tcgagggacc taataacttc gtatagcata cattatacga agttatatta agatcctcta 1380
170 gagtcgacct gcaggcatgc aagcttcagg gttgagatgt gtataagaga cagctgcatt 1440
172 aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct tccgcttctt 1500
```

RAW SEQUENCE LISTING

DATE: 04/28/2006

PATENT APPLICATION: US/10/505,328A

TIME: 09:38:57

Input Set : A:\Sequence.txt

Output Set: N:\CRF4\04282006\J505328A.raw

174 cgctcactga c 1511
177 <210> SEQ ID NO: 3
178 <211> LENGTH: 19
179 <212> TYPE: DNA
180 <213> ORGANISM: Artificial Sequence
182 <220> FEATURE:
183 <223> OTHER INFORMATION: chemically synthesized OE sequence
186 <400> SEQUENCE: 3
187 ctgtctctta tacacatct 19
190 <210> SEQ ID NO: 4
191 <211> LENGTH: 34
192 <212> TYPE: DNA
193 <213> ORGANISM: Artificial Sequence
195 <220> FEATURE:
196 <223> OTHER INFORMATION: chemically synthesized loxP site
199 <400> SEQUENCE: 4
200 ataacttcgt atagcataca ttatacgaag ttat 34
203 <210> SEQ ID NO: 5
204 <211> LENGTH: 996
205 <212> TYPE: DNA
206 <213> ORGANISM: Artificial Sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: chemically synthesized KmR gene
212 <400> SEQUENCE: 5
213 gcaaaaatta aaaatgaagt tttagcggta tcgaacccca gagtcccgct cagaagaact 60
215 cgtcaagaag gogatagaag gcgatgcgct gcgaatcggg agcggcgata ccgtaaagca 120
217 cgaggaagcg gtcagcccat tcgcgcgcaa gctcttcagc aatatcacgg gtagccaacg 180
219 ctatgtectg atagcgggtcc gccacaccca gccggccaca gtcgatgaat ccagaaaagc 240
221 ggccattttc caccatgata ttcggaagc aggcacgcc atgggtcacg acgagatcct 300
223 cgcgcgtcggg catccgcgcc ttgagcctgg cgaacagttc ggctggcgcg agccctgat 360
225 gctcttcgtc cagatcatcc tgatcgacaa gaccggcttc catccgagta cgtgctcgt 420
227 cgatgcgatg ttctgcttgg tggtcgaatg ggcaggtagc cggatcaagc gtatgcagcc 480
229 gccgcattgc atcagccatg atggatactt tctcggcagg agcaaggtga gatgacagga 540
231 gatcctgccc cggcacttgc cccaatagca gccagtcctt toccgcttca gtgacaacgt 600
233 cgagcacagc tgcgcaagga acgcccgtcg tggccagcca cगतagccgc gctgctcgt 660
235 cttggagttc attcagggca ccggacaggt cggctctgac aaaaagaacc gggcgccct 720
237 gcgctgacag ccggaacacg gcggcatcag agcagccgat tgtctgttgt gccagtcct 780
239 agccgaatag cctctccacc caagcggcg gagaaacctg gtgcaatcca tcttgttcaa 840
241 tcatgagaaa cgatcctcat cctgtctctt gatccactag attattgaag catttatcag 900
243 ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg 960
245 gttccggcca catttccccg aaaagtgcca cctgca 996
248 <210> SEQ ID NO: 6
249 <211> LENGTH: 947
250 <212> TYPE: DNA
251 <213> ORGANISM: Artificial Sequence
253 <220> FEATURE:
254 <223> OTHER INFORMATION: chemically synthesized GFP gene
257 <400> SEQUENCE: 6
258 attatttgta gagtcatcc atgccatgtg taatcccagc agcagttaca aactcaagaa 60

RAW SEQUENCE LISTING

DATE: 04/28/2006

PATENT APPLICATION: US/10/505,328A

TIME: 09:38:57

Input Set : A:\Sequence.txt

Output Set: N:\CRF4\04282006\J505328A.raw

```
260 ggaccatgtg gtcacgcttt tcgttgggat ctttcgaaag ggcagattgt gtcgacaggt 120
262 aatgggttgc tggtaaaagg acagggccat cgccaattgg agtattttgt tgataatggg 180
264 ctgctagtgt aacggatcca tcttcaatgt tgtggcgaat ttgaagtta gctttgattc 240
266 cattcttttg ttgtctgtcc gtgatgtata cattgtgtga gttatagtgt tactcgagtt 300
268 tgtgtccgag aatgtttcca tcttctttaa aatcaatacc ttttaactcg atacgattaa 360
270 caaggggtatc accttcaaac ttgacttcag cagcgtcttt gtagttcccg tcatctttga 420
272 aagatatagt gcgttcctgt acataacctt cgggcatggc actcttgaaa aagtcatggc 480
274 gtttcatatg atccggataa cgggaaaagc attgaacacc ataagagaaa gtagtgacaa 540
276 gtgttggcca tggaaacagg agttttccag tagtgcaaat aaatttaagg gtaagttttc 600
278 cgtatgttgc atcaccttca cctctccac tgacagaaaa ttgttgccca ttaacatcac 660
280 catctaattc aacaagaatt gggacaactc cagtgaagaag tcttctctct ttaactcatt 720
282 tttctaccgg taccggggga tectctagag tcgacctgca ggcattgcaag ctggcgtaa 780
284 tcatgggtcat agctgttttc tgtgtgaat tgttatccgc tcacaattcc acacaacata 840
286 cgagccggaa gcataaagtg taaagcctgg ggtgcctaag gtagtagcta actcacatta 900
288 attgcgttgc gtcactgccc cgtttccag tcgggaaatc caagggc 947
291 <210> SEQ ID NO: 7
292 <211> LENGTH: 1069
293 <212> TYPE: DNA
294 <213> ORGANISM: Artificial Sequence
296 <220> FEATURE:
297 <223> OTHER INFORMATION: chemically synthesized CmR gene
300 <400> SEQUENCE: 7
301 gcaaaaatta aaaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttgggtctg 60
303 acagttacca atgcttaatc agtgaggcac caataactgc cttaaaaaaa ttacgccccg 120
305 cectgccact catcgagta ctgttgtaat tcattaagca ttctgccgac atggaagcca 180
307 tcacagacgg catgatgaac ctgaatcgcc agcggcatca gcacctgtc gccttgcgta 240
309 taatatttgc ccattggtgaa aacggggggcg aagaagtgtt ccatattggc cactgtttaa 300
311 tcaaaactgg tgaacctcac ccagggattg gctgagacga aaaacatatt ctcaataaac 360
313 cctttaggga aataggccag gttttcaccg taacacgcca catottgcga atatattgtg 420
315 agaaactgac ggaatcgtc gtggtattca ctccagagcg atgaaaacgt ttcagtttgc 480
317 tcatggaaaa cgggtgtaaca aggggtgaaca ctatcccata toaccagctc accgtcttcc 540
319 attgccatac ggaatttcgg atgagcattc atcaggcggg caagaatgtg aataaaggcc 600
321 ggataaaaact tgtgcttatt tttctttacg gtctttaaaa aggcctgaat atccagctga 660
323 acggctctgg ttaggttaca ttgagcaact gactgaaatg cctcaaaatg ttctttacga 720
325 tggcatttgg atatatcaac ggtggtatat ccagtgtatt tttttcccat tttagcttcc 780
327 ttagctcctg aaaaatctga taactcaaaa aatacggccg gtagtgatct tatttcatta 840
329 tggtgaaagt tggaaacctc tacgtgcccga tcaacgtctc attttcgcca aaagtgggoc 900
331 cagggcttcc cggtatcaac agggacacca ggatttatat attctcgcaa gtgatcttcc 960
333 gtcacaggta tttattcggc gcaaaagtgc tcgggtgatg ctgccaaact actgatttag 1020
335 tgtatgatgg tgtttttgag gtgctccagt ggcttctgtt tcatcagc 1069
338 <210> SEQ ID NO: 8
339 <211> LENGTH: 19
340 <212> TYPE: DNA
341 <213> ORGANISM: Artificial Sequence
343 <220> FEATURE:
344 <223> OTHER INFORMATION: chemically synthesized primer-pMODFP-1
347 <400> SEQUENCE: 8
348 attcaggctg cgcaactgt 19
351 <210> SEQ ID NO: 9
```

RAW SEQUENCE LISTING

DATE: 04/28/2006

PATENT APPLICATION: US/10/505,328A

TIME: 09:38:57

Input Set : A:\Sequence.txt

Output Set: N:\CRF4\04282006\J505328A.raw

352 <211> LENGTH: 22
353 <212> TYPE: DNA
354 <213> ORGANISM: Artificial Sequence
356 <220> FEATURE:
357 <223> OTHER INFORMATION: chemically synthesized primer-pMODRP-1
360 <400> SEQUENCE: 9
361 tcagtgagcg aggaagcgga ag 22
364 <210> SEQ ID NO: 10
365 <211> LENGTH: 28
366 <212> TYPE: DNA
367 <213> ORGANISM: Artificial Sequence
369 <220> FEATURE:
370 <223> OTHER INFORMATION: chemically synthesized primer-Tn5Ext
373 <400> SEQUENCE: 10
374 agcatatcatt atacgaagtt atattaag 28
377 <210> SEQ ID NO: 11
378 <211> LENGTH: 35
379 <212> TYPE: DNA
380 <213> ORGANISM: Artificial Sequence
382 <220> FEATURE:
383 <223> OTHER INFORMATION: chemically synthesized primer-Arb1
386 <400> SEQUENCE: 11
W--> 387 ttgagcgata gacgtacgat nnnnnnnnnn gatat
390 <210> SEQ ID NO: 12
391 <211> LENGTH: 20
392 <212> TYPE: DNA
393 <213> ORGANISM: Artificial Sequence
395 <220> FEATURE:
396 <223> OTHER INFORMATION: chemically synthesized primer-Arb2
399 <400> SEQUENCE: 12
400 ttgagcgata gacgtacgat
403 <210> SEQ ID NO: 13
404 <211> LENGTH: 25
405 <212> TYPE: DNA
406 <213> ORGANISM: Artificial Sequence
408 <220> FEATURE:
409 <223> OTHER INFORMATION: chemically synthesized primer-Tn5Int
412 <400> SEQUENCE: 13
413 tcgacctgca ggcattgcaag cttca 25

pls explain "N" locations. 35

See error explanation on page 7.

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 04/28/2006
PATENT APPLICATION: US/10/505,328A TIME: 09:38:58

Input Set : A:\Sequence.txt
Output Set: N:\CRF4\04282006\J505328A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:11; N Pos. 21,22,23,24,25,26,27,28,29,30

VARIABLE LOCATION SUMMARY

DATE: 04/28/2006

PATENT APPLICATION: US/10/505,328A

TIME: 09:38:58

Input Set : A:\Sequence.txt

Output Set: N:\CRF4\04282006\J505328A\raw

Use of n's or Xaa's (NEW RULES):

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of <220> to <223> is MANDATORY if n's or Xaa's are present.

in <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

Seq#:11; N Pos. 21,22,23,24,25,26,27,28,29,30 ✓

ERROR Explanation: 2

VERIFICATION SUMMARY

DATE: 04/28/2006

PATENT APPLICATION: US/10/505,328A

TIME: 09:38:58

Input Set : A:\Sequence.txt

Output Set: N:\CRF4\04282006\J505328A.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:387 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:11
L:387 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:11
L:387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0

file://C:\CRF4\Outhold\VsJ505328A.htm

4/28/2006